

БАЗЫ ДАННЫХ ОБЛАЧНЫХ ТЕХНОЛОГИЙ УПРАВЛЕНИЯ

Погребняк Б.И., Харьковский национальный университет городского хозяйства имени А. Н. Бекетова

Базы данных (БД) – это «рабочие лошадки» информационных систем управления. Работа с любой из таких систем по существу является взаимодействием с информацией, которая хранится в ее базе данных. Наблюдаемый же сегодня повсеместный переход от традиционных систем управления к их облачным аналогам выдвигает целый ряд специфических требований к системам управления такими базами данных (СУБД).

Количество имеющихся на рынке СУБД растет лавинообразно, и сравнивать одну с другой становится все труднее. К счастью, большинство из них попадает в одну из нескольких категорий. При этом наиболее существенным признаком деления на такие категории является модель данных, реализованная в конкретной СУБД.

Понятие модели данных является центральным понятием в области баз данных [1]. Под ней понимается набор принципов, определяющих организацию логической структуры хранения данных и совокупность операций над ними. Наиболее зрелые решения демонстрируют СУБД основанные на реляционной модели данных такие, например, как Oracle, MySQL или PostgreSQL [2]. Впервые реляционная модель данных была предложена доктором Эдгаром Коддом (Edgar Codd), сотрудником компании IBM, в 1970 году.

В этой модели все данные представлены в виде простых таблиц, состоящих из строк и столбцов, на пересечении которых расположены конкретные значения. При этом столбцы определяют структуру таблицы, а строки – число записей в ней. Как правило, одна база данных содержит несколько таблиц, которые могут быть, как связаны друг с другом, так и независимы друг от друга. С целью удаления в такой БД избыточности и непротиворечивости над ней необходимо производить специальные процедуры нормализации, которые заключаются в разнесении данных по нескольким таблицам. Среди прочего, нормализованная БД характеризуется тем, что каждый элемент данных хранится только в одном месте.

Все операции над данными в реляционной СУБД сводятся к операциям над таблицами. При этом операции можно группировать в так называемые транзакции, выполняемые по принципу «все или ничего». Язык, при помощи которого описываются такие манипуляции, называется SQL (Structured Query Language – язык структурированных запросов). Он является стандартом de-facto языка работы с реляционными базами данных.

Производители большинства реляционных СУБД всегда пытались позиционировать свои системы, как решения «все в одном». И они действительно идеально подходят, когда необходимо обеспечить высокую надежность хранения информации, например, в банковских и финансовых приложениях. Однако при переходе к распределенным, высоконагруженным и масштабируемым Web-приложениям оказалось, что реляционная модель данных не является универ-

сальной, и у нее есть свои слабые места. Это обусловлено тем, что изначально она проектировалась для реализации в монолитном приложении, выполняющемся на локальном компьютере с ограниченными вычислительными ресурсами. Компьютерные сети тогда были в самом зачаточном состоянии. В связи с этим в течение нескольких последних лет все большую популярность приобретают так называемые NoSQL-технологии. Собираемый термин NoSQL расшифровывается как Not only SQL (не только SQL). То есть, NoSQL – это только название для целого спектра самых различных решений. Общим у них является только одно – альтернатива реляционному подходу хранения и манипулирования данными. NoSQL стремится к тому, чтобы для определенной задачи был выбран такой инструмент, который бы решал эту задачу наилучшим образом.

Простейшие NoSQL-хранилища данных (такие, например, как Memcached, Cassandra или Redis), построенные по принципу «ключ/значение», обладают наибольшей производительностью при наименьшей функциональности. На противоположной стороне этого спектра СУБД располагаются традиционные реляционные решения с наибольшей функциональностью «из коробки» и наименьшей производительностью. А одним из наиболее интересных и перспективных решений в этом диапазоне является документно-ориентированная СУБД MongoDB [3, 4].

MongoDB – это система управления базами данных, специально «заточенная» под Web-приложения и инфраструктуру Интернета. На самом вершине информационной модели данных MongoDB находится объект базы данных. Он состоит из коллекций, которые, в свою очередь, состоят из отдельных документов. Каждый документ в коллекции представляет собой набор свойств, состоящий из названий и их значений. Значения свойств могут быть представлены простыми типами данных, такими как строка, число или дата. Но могут быть также массивами или даже другими документами. То есть, в документно-ориентированной модели данных MongoDB каждый документ может иметь свою структуру. На входе они представляются в формате JSON (JavaScript Object Notation), т. е. «родным» языком для MongoDB является JavaScript. Самый актуальный перечень промышленных внедрений MongoDB представлен в [5].

Литература

1. Дейт К. Введение в системы баз данных. – 8-е изд. – М.: Вильямс, 2006. – 1328 с.
2. MySQL 5/ М.В. Кузнецов, И.В. Симдянов. – СПб.; БХВ-Петербург, 2010. – 1024 с.: ил.
3. The Little MongoDB Book By Karl Seguin – [Электронный ресурс]. – Режим доступа: <https://github.com/karlseguin/the-little-mongodb-book>.
4. Бэнкер К. MongoDB в действии. / Пер. с англ. Слинкина А. А. – М.: ДМК Пресс, 2012. – 394 с: ил.
5. Актуальный перечень промышленных внедрений MongoDB – [Электронный ресурс]. – Режим доступа: <http://mng.bz/z2CH>.